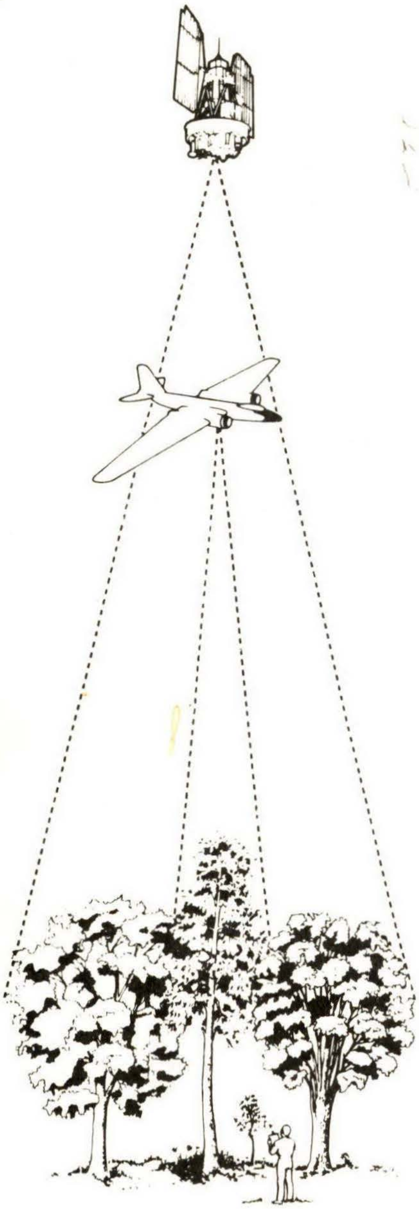
NATIONWIDE FORESTRY APPLICATIONS PROGRAM



TEN-ECOSYSTEM STUDY (TES) SITE IX, WASHINGTON COUNTY, MISSOURI, FINAL REPORT

W. H. Echert
Lockheed Electronics Company, Inc.
Systems and Services Division
Houston, Texas 77058

Contract NAS 9-15800
JSC-14657
LEC-13000
March 1979

Prepared for
EARTH OBSERVATIONS DIVISION
SPACE AND LIFE SCIENCES DIRECTORATE
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS 77058

NASA
National Aeronautics and
Space Administration

Lyndon B. Johnson Space Center
Houston, Texas 77058

NATIONAL FOREST SYSTEM · COOPERATIVE FORESTRY · FORESTRY RESEARCH

FOREST SERVICE
U.S. Department of Agriculture

NOTE: In 1976, the Nationwide Forestry Applications Program was expanded from a Regional project by co-operative agreement between the Forest Service, U. S. Department of Agriculture, and the National Aeronautics and Space Administration (NASA). The Program is designed to sponsor research and development on the application of remote sensing analysis techniques to problems arising from the need to inventory, monitor and manage forests and rangelands, including the assessment of impacts on forest stands from insect and disease damage.

1. Report No. JSC-14657	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Ten-Ecosystem Study (TES) Site IX, Washington County, Missouri, Final Report		5. Report Date March 1979
		6. Performing Organization Code
7. Author(s) W. H. Echert Lockheed Electronics Company, Inc.		8. Performing Organization Report No. LEC-13000
9. Performing Organization Name and Address Lockheed Electronics Company, Inc. Systems and Services Division 1830 NASA Road 1 Houston, Texas 77058		10. Work Unit No.
		11. Contract or Grant No. NAS 9-15800
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Lyndon B. Johnson Space Center Houston, Texas 77058 Technical Monitor: R. E. Joosten		13. Type of Report and Period Covered Type II Final Report
		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

This final report presents the results, conclusions, and recommendations obtained from the automatic data processing analysis of Landsat and photographic data acquired over Washington County, Missouri, Site IX of the Ten-Ecosystem Study. This site was selected to be representative of the Central Hardwood Ecosystem.

The analysis of the processing results has led to the following conclusions:

1. Sufficient spectral separability exists among softwood, hardwood, grassland, and water to develop a Level II classification and inventory.

2. Using the tested automatic data processing technology, softwood and grassland signatures can be extended across the county with acceptable accuracy; with more dense sampling, the hardwood signature probably could also be so extended.

3. The best season for mapping this ecosystem using Landsat data is early fall.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Forestry Grassland Inventory Landsat Remote sensing	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	85	

*For sale by the National Technical Information Service, Springfield, Virginia 22161

JSC Form 1424 (Rev Nov 75) NASA — JSC

PREFACE

To prepare for future nationwide forest and grass renewable resource inventories using automatic data processing and remote sensing technology, the National Aeronautics and Space Administration at the Lyndon B. Johnson Space Center and the Forest Service of the U.S. Department of Agriculture divided the continental United States into 10 generalized forest and grassland ecosystems to conduct a study known as the Ten-Ecosystem Study. The Ten-Ecosystem Study, a part of the Nationwide Forestry Applications Program, used Landsat data, supporting aircraft imagery, and ancillary information to make a forest, grassland, and inland water area inventory of chosen sites* within the 10 ecosystems.

The primary objectives of the Ten-Ecosystem Study were to

1. Investigate the feasibility of using automatic processing of remotely sensed data to inventory forest, grassland, and inland water areas within administrative boundaries for specified ecosystems of the United States
2. Identify automatic data processing analysis problems related to each site or ecosystem and recommend solutions
3. Define the requirements for an automatic data processing system to make a nationwide forest and grassland inventory

Secondary objectives of the Ten-Ecosystem Study were to

1. Determine the accuracy of mapping Level II features (softwood, hardwood, grassland, and water) using computer-aided classification

*Only nine different test sites were required because one site was selected to represent two ecosystems.

2. Establish the best season for accurately mapping each site or ecosystem

3. Provide the Forest Service with findings and conduct an evaluation workshop to exchange ideas and receive Forest Service feedback

The Ten-Ecosystem Study consisted of two parts:

1. A type separability study to determine the range of possibilities for mapping forest, grassland, and inland water details and the corresponding mapping accuracies obtainable by automatic processing analysis of remotely sensed data

2. A simulated inventory study to determine how successfully automatic data processing technology can extend limited ground truth to make large-area inventories

This document is the final of four reports covering the study conducted at the Washington County, Missouri, site. Based on the conditions existing at the time of this study and on the procedures used, the results indicate that softwood, hardwood, and grassland can be classified with an estimated accuracy of 85 percent ± 6.2 percent.

This report was prepared by Lockheed Electronics Company, Inc., under Contract NAS 9-15800, Job Order 75-325, Action Document 63-1557-5325-32. Distribution of this report has been approved by the supervisor of the Forestry Applications Section and the manager of the Earth Observations Exploratory Studies Department.

Numerous individuals participated in the analysis of the Washington County, Missouri, site. B. F. Edwards, Principal Scientist, developed the regression statistics. L. R. Hall, Senior Associate Scientist, contributed technical support and

assistance in processing Landsat data via the Interactive Multispectral Image Analysis System, Model 100, and the Data Analysis Station of the Passive Microwave Imaging System. Scientists J. F. Ward and R. H. Almond provided support during the site familiarization trip. Scientists D. R. King and W. H. Parkhurst acquired source materials and evaluated imagery during the preliminary analysis.

CONTENTS

TABLES

FIGURES

ABBREVIATIONS

ADP	automatic data processing
CIR	color infrared
DAS	Data Analysis Station
ERIPS	Earth Resources Interactive Processing System
Image 100	Interactive Multispectral Image Analysis System, Model 100
MSS	multispectral scanner
NASA	National Aeronautics and Space Administration
PCC	percent correct classification
pixel	picture element
PMIS	Passive Microwave Imaging System
PDP 11/45	Programmed Data Processor, Model 11/45
PSU	primary sampling unit; first stage of a two-stage sampling scheme (It consists of an area 50 pixels by 50 pixels in size located at random throughout the site.)
SSDA	Sequential Similarity Detection Algorithm
SSU	secondary sampling unit; second stage of a two-stage sampling scheme (It consists of a 2-pixel by 2-pixel block located at random within each PSU.)
TES	Ten-Ecosystem Study
USDA	U.S. Department of Agriculture

1. INTRODUCTION

As part of the Ten-Ecosystem Study (TES), Washington County, Missouri (Site IX), was selected to represent the Central Hardwood Ecosystem (ref. 1). Two studies, a type separability and a simulated inventory, were made to determine the feasibility of using remote sensing technology to inventory forest, grassland, and inland water areas. The type separability study was done to determine the maximum level of detail and corresponding mapping accuracies obtainable using automatic processing analysis of remotely sensed data. In this study, ground truth and aerial photographs were used to select training fields for the computer-aided classification of the county into softwood, hardwood, grassland, and water. The simulated inventory study was conducted to determine how successfully automatic data processing (ADP) technology can extend limited ground truth to make large-area inventories. In this study, only aerial photographs were used to select training fields. These training fields, located in a specified 10 percent of the study site, were used to classify the entire county into softwood, hardwood, and grassland. (Water was not classified because the simulated test area did not contain water bodies extensive enough to be resolved by the Landsat sensor.)

1.1 SCOPE

In the TES, Landsat multispectral scanner (MSS) data were used as the mapping data base. The National Aeronautics and Space Administration (NASA) high-altitude color-infrared (CIR) aerial photographs (1:120 000 scale) were used as the photointerpretation base. To aid in the analysis, TES personnel used remote sensing application publications, research reports, and personal knowledge obtained from Forest Service contacts in addition to an onsite survey.

Each site was to cover one county or 360 000 square hectometers (889 579 acres), whichever was smaller. Two Landsat data sets representing different seasons were used to determine the effects of single versus temporal data sets. Classification maps of the study site were generated and statistically evaluated, and estimated proportions were compared to county inventory figures.

1.2 ANALYSIS LEVELS

The TES processing procedures (ref. 2) consisted of two distinct steps: (1) a type separability study and (2) a simulated inventory study. Both procedures followed a supervised method for determining class signatures. The type separability study was designed to determine the separability of Level II features (softwood, hardwood, grassland, and water) before executing the simulated inventory study. This separability was to be determined by using ground-verified training fields, obtaining their corresponding Landsat signatures, classifying the area, and computing the percent correct classification (PCC) based on the number of correctly classified picture elements (pixels) in each training field. This PCC would thus indicate the consistency of the signatures of the different classes. If the PCC results of the separability study met or exceeded the accuracy requirements for a Level II classification, a Level III separability study was to be undertaken. This, of course, would have required the use of Level III training fields; that is, the separation of hardwoods and softwoods into species such as white oak, hickory, pine, and cedar. However, during the field survey, such a separation was proven to be infeasible.

The second step of the processing task was the simulated inventory study. This study also followed a supervised classification method; however, in this case, the location of each

training field was restricted to a predesignated area which encompassed approximately 10 percent of the site. Another constraint was that, in locating and identifying training fields, the image analyst was restricted to the use of aerial photographs and denied recourse to onsite verification. The purpose of the simulated inventory was to determine the accuracy of a supervised classification using training fields based on limited information and selected from only a small portion of the site. In essence, this process is signature extension, as defined by reference 2.

1.3 ADP EQUIPMENT

As an ADP study using Landsat data, TES was designed around an interactive computer system that allows the analyst to interject and modify input parameters in near real time, enhance the image display, and classify the data as desired. The General Electric Interactive Multispectral Image Analysis System, Model 100 (Image 100) was selected as the main computer system. The primary function of the Image 100 as used in the TES study was to extract thematic information from Landsat data tapes. The system performs thematic information extraction in a supervised mode; i.e., the user defines a small portion of an area of interest and the Image 100 classifies the remainder of the area by comparing the Landsat spectral properties of the defined area (training field) to those of the total area displayed on the cathode-ray tube. The user can modify the input parameters before executing the extraction process, observe the results, and, if necessary, continue to modify the parameters until the desired results are obtained. Additionally, the Earth Resources Interactive Processing System (ERIPS) and the Data Analysis Station (DAS) of the Passive Microwave Imaging System (PMIS) supported the TES operational system, as did the Dell Foster coordinatograph.

2. SITE DESCRIPTION

Washington County, Missouri, was selected as representative of the Central Hardwood Region (fig. 2-1) and designated TES Site IX. The test site is located in the mideastern portion of the state, approximately 80 kilometers (50 miles) south-southwest of St. Louis. The central geographic coordinates of the site are 37°57' north and 95°59' west. The site includes the entire county, extending east to west 40.1 kilometers (24.9 miles) and north to south 51.8 kilometers (32.2 miles); it encompasses a total land area of some 196 843 square hectometers (486 400 acres). The town of Potosi, with a population of 2761, is the county seat and the largest urban development within the county, the total population for the county being 15 015 (ref. 3).

Although Washington County is the center of the hardwood-veneer industry in Missouri, it is more noted for its mining industry and has often been referred to as the "Barite Capital of the World." As a result, the hardwoods of the region have been in demand for railroad ties and mine-support timbers as well as for hardwood flooring and furniture veneers. A summary of the county forest statistics is given in table 2-1 (ref. 4).

Approximately 30 percent of the land in Washington County is a part of the Mark Twain National Forest.[1]

2.1 GEOLOGY

Washington County is situated on the northeastern flank of the Ozark uplift, which extends throughout much of the southern part

---

[1]This forest was formerly called the Clark National Forest, but, on February 17, 1976, President Ford signed an order changing the name to Mark Twain National Forest (ref. 5).

of Missouri and infringes on the borders of Kansas, Oklahoma, and Arkansas. The uplift is a very broad, gentle structural swell or dome and forms an irregular belt approximately 241 kilometers (150 miles) wide and 402 kilometers (250 miles) long, with the tectonic forces appearing to be concentrated in the St. Francois Mountains. These mountains are composed of Precambrian granites, rhyolites, and other crystalline rocks, reaching an elevation of 518 meters (1700 feet). Throughout the Ozark region, the younger Cambrian strata unconformably overlie the Precambrian core (see fig. 2-2). The strata dip gently to the northwest and south and become increasingly inclined in the east and southeast directions. However, in much of Missouri and especially in Washington County, the inclination of the dipping beds is slight, with the regional dip generally disguised by minor local folding. The Ozarks are now represented as a series of ancient rocks encircled by younger strata that dip gently away from the central core (ref. 6).

2.2 TOPOGRAPHY

Washington County rests on a major topographical feature known as the Ozark Plateau, which contains three major landforms — the Springfield Plateau, the Salem Upland, and the St. Francois Mountains. Washington County is in the Salem Upland, which occupies the east-central part of Missouri. Although lower in elevation than the Springfield Plateau, the Salem Upland is extremely dissected, and the topographic expression in the western portion of the county is generally in the form of steep ridges with narrow and poorly defined valleys. The elevations range from 213 meters (700 feet) to 335 meters (1100 feet). To the east, the valleys become broader and the ridges lose their sharp definition, giving rise to rolling hills and more open

country. The mean elevation of the ridges is approximately 305 meters (1000 feet), while the highest point in the county (Little Pilot Knob) has an elevation of 430 meters (1412 feet).

The drainage follows a modified dendritic pattern, becoming more angular to the northwest, where the drainage may be an expression of the underlying Ordovician facies and/or local faulting as compared to the drainage to the south and east which is underlain by the Upper Cambrian series. The area is dissected by a few well-developed streams with numerous intermittent tributaries.

2.3 CLIMATE

Missouri has a typical continental climate characterized by frequent and sometimes extreme changes in weather. The summers can be hot and humid, whereas the winters may have periods of extreme cold. The temperature exceeds 38° C (100° F) in 7 out of every 10 years, and such high temperatures normally last about 3 days. In 6 of 10 years, the temperature will drop below -18° C (0° F); however, this condition seldom lasts more than a few days.

The average length of the growing season varies from 175 to 199 days, whereas precipitation seldom varies from an average range of 102 to 114 centimeters (40 to 45 inches) per year. May is the wettest month, averaging more than 13 centimeters (5 inches) of rainfall; and October is the driest month, with about 6 centimeters (2.5 inches) of rainfall on the average. Snowfall averages less than 18 centimeters (7 inches) per year, occurring three or four times annually and quickly melting (ref. 5).

2.4 SOILS

The most extensive soils in the county are those formed by the weathering of the underlying carbonate/cherty formations; the

Clarksville-Fullerton-Talbott soils extend over 80 percent of the area. The Hagerstown-Tilsit and Ashe-Tilsit-Hagerstown soil associations found in the eastern and southeastern parts of the county make up the balance of the soils. In addition, there are alluvial deposits along the streams and colluvial deposits on the lower slopes. Soils on the upper slopes generally are infertile and stony with chert fragments on or near the surface (ref. 8).

2.5 WATER RESOURCES

The water resources of the national forests in Missouri are sufficient for all anticipated uses through the year 1990 (ref. 5). Within the boundaries of Washington County are numerous small tailing ponds formed by surface mining. The only sizable lake is Sunnen Lake located on the Fourche a Renault and consisting of approximately 81 square hectometers (200 acres).

The major drainage of the area is modified dendritic (fig. 2-3); the Mineral Fork and Fourche a Renault drain the area to the northeast, and Lost, Hazel, and Cub Creeks drain the area to the west. The major drainage system in the southeast is formed by the Big River and its tributaries. The streams are generally unpolluted and free of turbidity.

2.6 VEGETATION

The Washington County study site is completely contained in the Central Hardwood Ecosystem. As reference 7 indicates, this region contains five major forest types: oak-hickory, mixed hardwood, elm-ash-maple, pine-hardwood, and eastern red cedar. Of these five, the major type in Washington County is oak-hickory. White oak, black oak, post oak, and red oak make up the oak forest species. Other hardwoods of the elm-ash-maple type and

occasionally cottonwood and sycamore are found in minor stands on stream-cut benches.

The major softwood species is the shortleaf pine, which occurs intermixed with hardwoods on the upper slopes and ridges. With the exception of plantation pines, the highest concentration of this species is in the west-central and southwest portions of the county. Throughout the rest of the area, the pine is well dispersed. The eastern red cedar is of secondary importance, occurring in small plantations (1 to 4 square hectometers or 2 to 10 acres), on the fringes of the hardwood stands, or in glades.

Because of increased commercial activity, silvicultural practices since 1934 have led to a greater inventory of shortleaf pine in the county. Early planting was confined to abandoned fields. Later planting was by blocks (up to two sections each) — mostly under the cover of hardwood canopies.

Viewed broadly, the vegetation of the region is a forest flora; the oak-hickory and oak-pine types of woodland dominate in the highlands, and willow, sycamore, and maple prevail in the bottomlands, with brush light to medium in density. Grassland, in the form of well-developed pastures, is confined to the valleys and terraces adjacent to the larger streams. Cropland is practically nonexistent except for small fields that support only the local populace.

TABLE 2-1.— WASHINGTON COUNTY FOREST STATISTICS

Category	Square hectometers ($\times 10^3$)	Acres ($\times 10^3$)
Commercial forest		
Softwood:[a]		
Shortleaf pine	4.17	10.3
Eastern red cedar	.45	1.1
Mixed softwood-hardwood:		
Shortleaf pine-oak	6.80	16.8
Eastern red cedar-hardwood	.85	2.1
Hardwood:		
Black oak and scarlet oak	67.06	165.7
White oak	30.60	75.6
Post oak and blackjack oak	23.92	59.1
Maple and beech	2.80	6.9
Elm, ash, and cottonwood	2.35	5.8
Oak, gum, and cypress	.16	.4
Nonstocked	4.53	11.2
Commercial forest total	143.67	355.0
Noncommercial forest	1.62	4.0
All forest types	145.29	359.0
Land area	196.84	486.4

[a]For the purpose of this report, softwood was combined with mixed softwood-hardwood for a total area of 12.27 square hectometers (30.3 acres), while the hardwood area included the noncommercial forests for a total of 133.02 square hectometers (328.7 acres). These figures approximate those tabulated in reference 7.

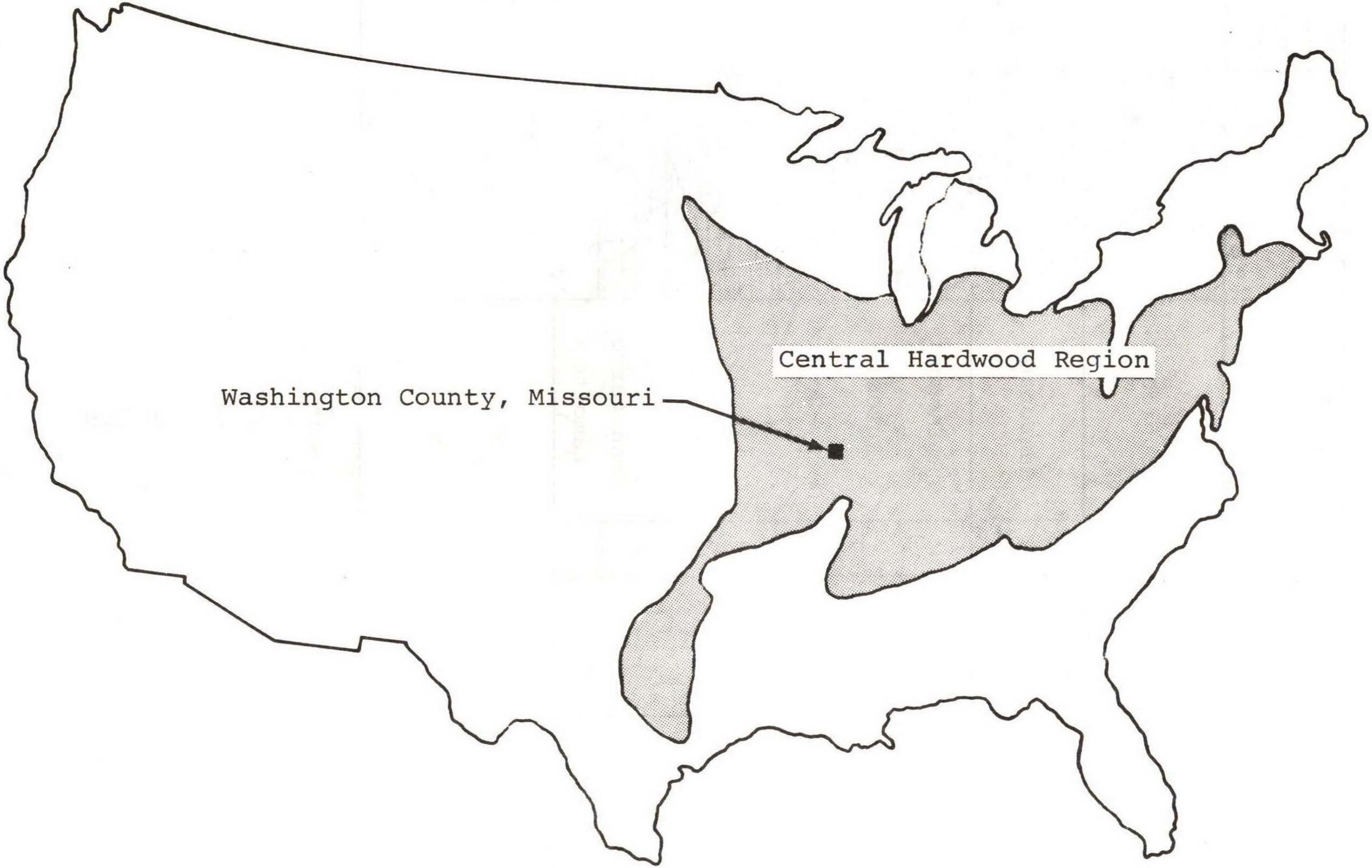


Figure 2-1.– Location of Central Hardwood Region and TES Site IX (Washington County, Missouri).

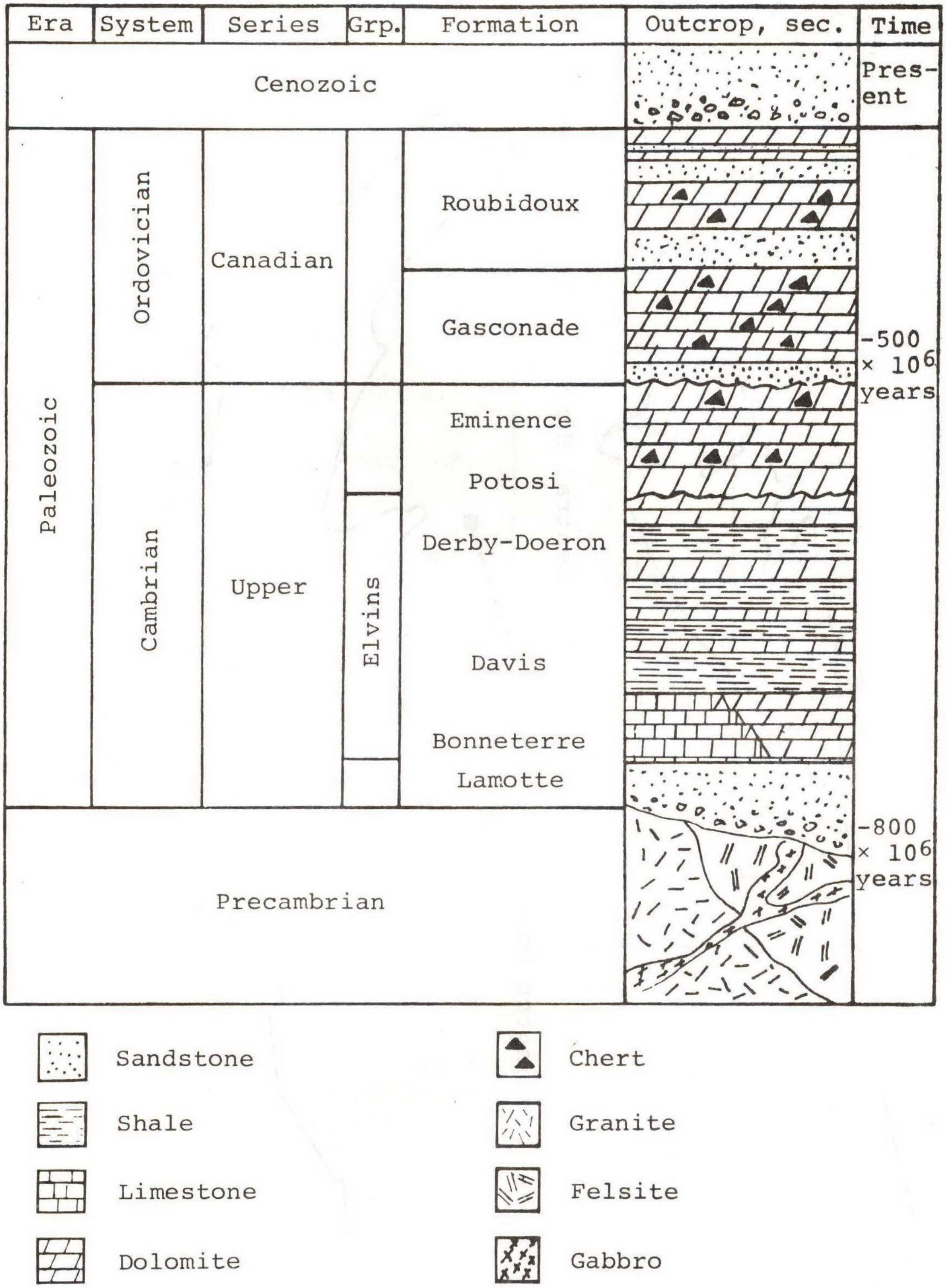


Figure 2-2.— Washington County stratigraphic column.

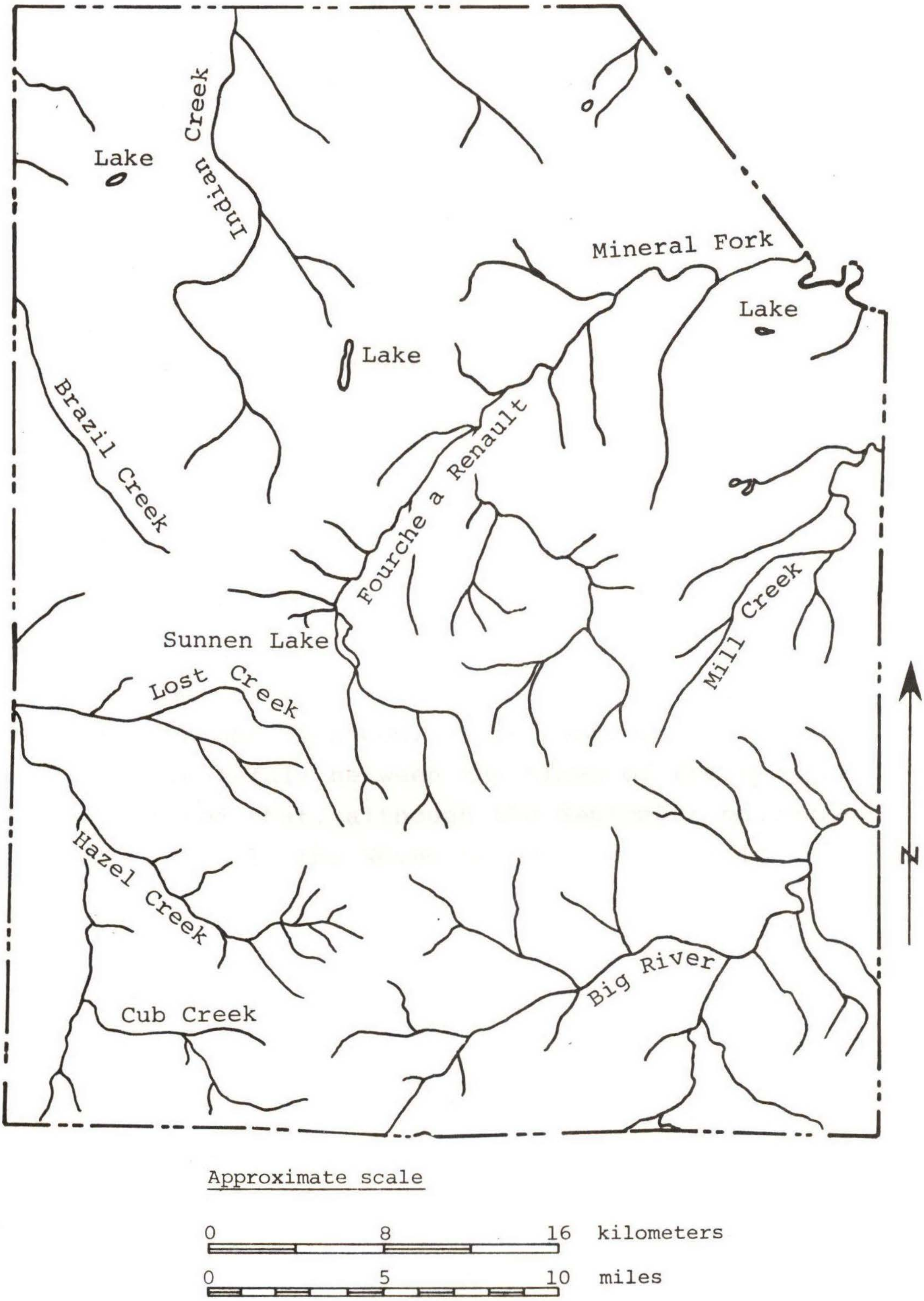


Figure 2-3.– Washington County drainage.

3. EQUIPMENT AND MATERIALS

3.1 AERIAL PHOTOGRAPHIC DATA

Aerial photography from two flights of Mission 289 (September and November 1974) was used. The CIR photography was collected by NASA WB-57 aircraft from an altitude of 18 293 meters (60 000 feet), using a 152-millimeter (6-inch) focal length lens to produce a nominal scale of 1:120 000. The two flights provided approximately 85-percent coverage of the test site, as is shown in figure 3-1.

The quality of the aerial photographs was generally good with the exception of vignetting around the periphery which inhibited stereoscopic observation in the overlapping areas. The 60-day separation between flights was sufficient to cause a decided change in the color hues and tones. The September photography exhibited a wide range of magenta and cyan, whereas the November imagery was much more subdued. This contrast in hue and tone was caused by leaf-fall between the times of the two flights. It should be noted that, although the September photographs provided more detail, the November photographs allowed more definitive separation between the pines and hardwoods.

3.2 LANDSAT IMAGERY

Selection of the Landsat imagery was based on established criteria which considered cloud cover, temporal separation, scanner noise, and obvious factors that could degrade the ADP procedure. On the basis of these criteria, eight Landsat scenes were obtained from the NASA Goddard Space Flight Center. These scenes were then evaluated to determine which data provided maximum forest information. To form the temporal data set, the two scenes that had the highest ratings and provided sufficient temporal separation were chosen. Of the eight

scenes evaluated, the two that most nearly fulfilled these requirements were Landsat scenes 1737-16025 (July 30, 1974) and 1845-15591 (November 15, 1974).

See appendix A for the interpretation and evaluation of the aerial photography and Landsat imagery.

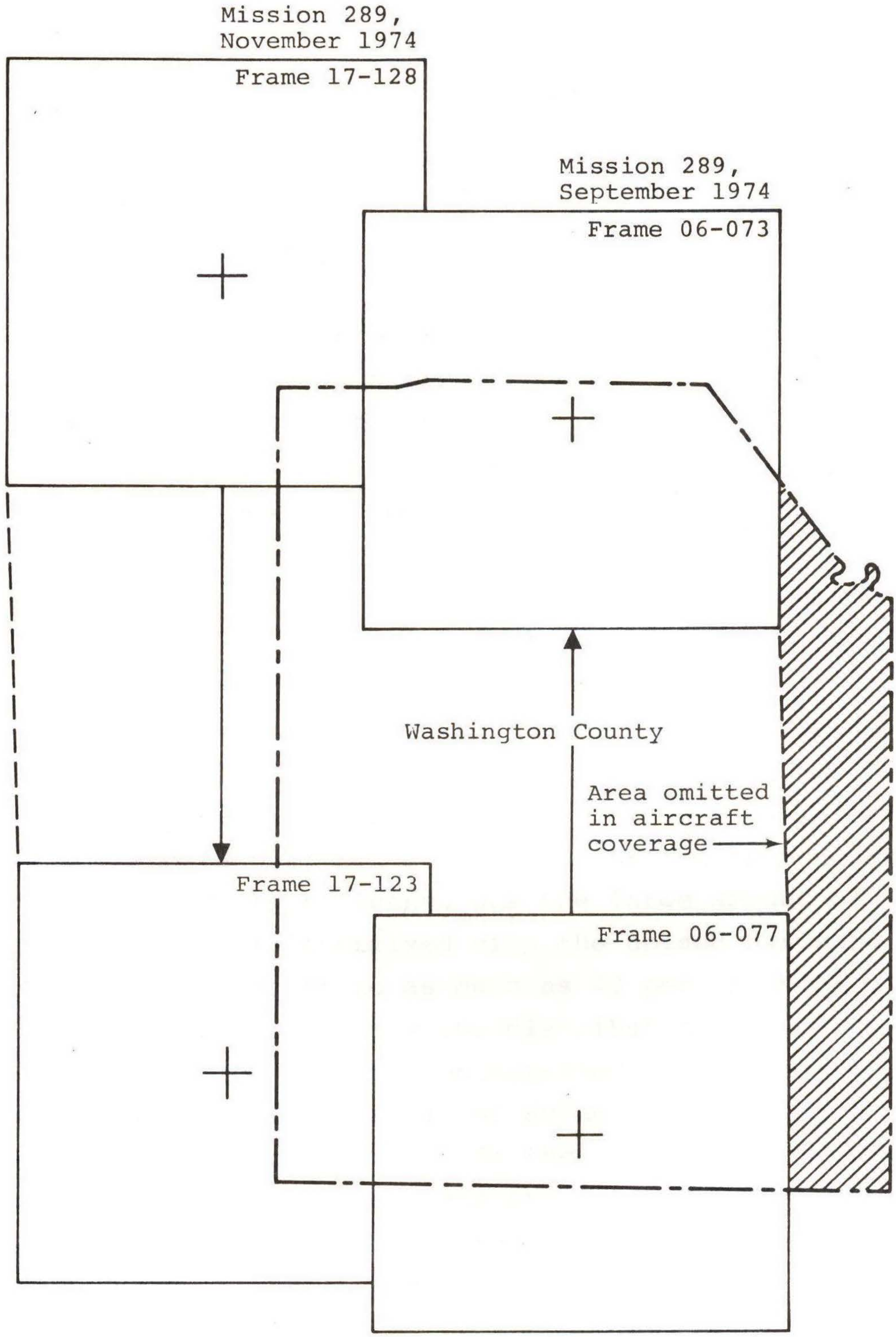


Figure 3-1.– Aerial photographic coverage of Washington County.

4. SITE FAMILIARIZATION TRIP

4.1 INTRODUCTION

The site familiarization trip to Washington County, Missouri, was conducted from October 31 through November 4, 1977. The survey group was made up of three Lockheed Electronics Company, Inc., personnel and a representative from NASA's Lyndon B. Johnson Space Center. The purpose of the trip was to familiarize the survey team with the various forest ecological units in Washington County and to correlate laboratory interpretations of aerial photographs with onsite observations of 122 preselected training fields. Plans included coordinating activities with the resident scientists from the Forest Service of the U.S. Department of Agriculture (USDA) in order to gain more practical knowledge of the area and to obtain ancillary data such as timber stand maps and supporting documentation.

4.2 RESULTS

The field survey of Washington County provided much information that was not apparent from the photointerpretation conducted in the laboratory. Of significance was the large amount of softwood (shortleaf pine) intermixed with the upland hardwood (the mix ranged from 5 percent to as much as 40 percent pine). Another significant factor was the distribution of lowland hardwoods, which were found to be concentrated on stream-cut benches. These stands consisted primarily of sycamores, maples, and willows, with thick underbrush. Of the hardwoods observed, stands of white oaks appeared to dominate the landscape and were generally intermixed with minor stands of black oaks, post oaks, and scarlet oaks in addition to hickories.

Some difficulty was encountered in correlating the September photographs (leaf-on condition) with ground observations made after the deciduous trees had lost their leaves. Subtle changes

that could be traced on the aerial photographs could not be traced on the ground.

The locations of the training fields and routes leading to them are shown in figure 4-1. Because of insufficient time in the field, only 100 of the 122 preselected training fields were visited.

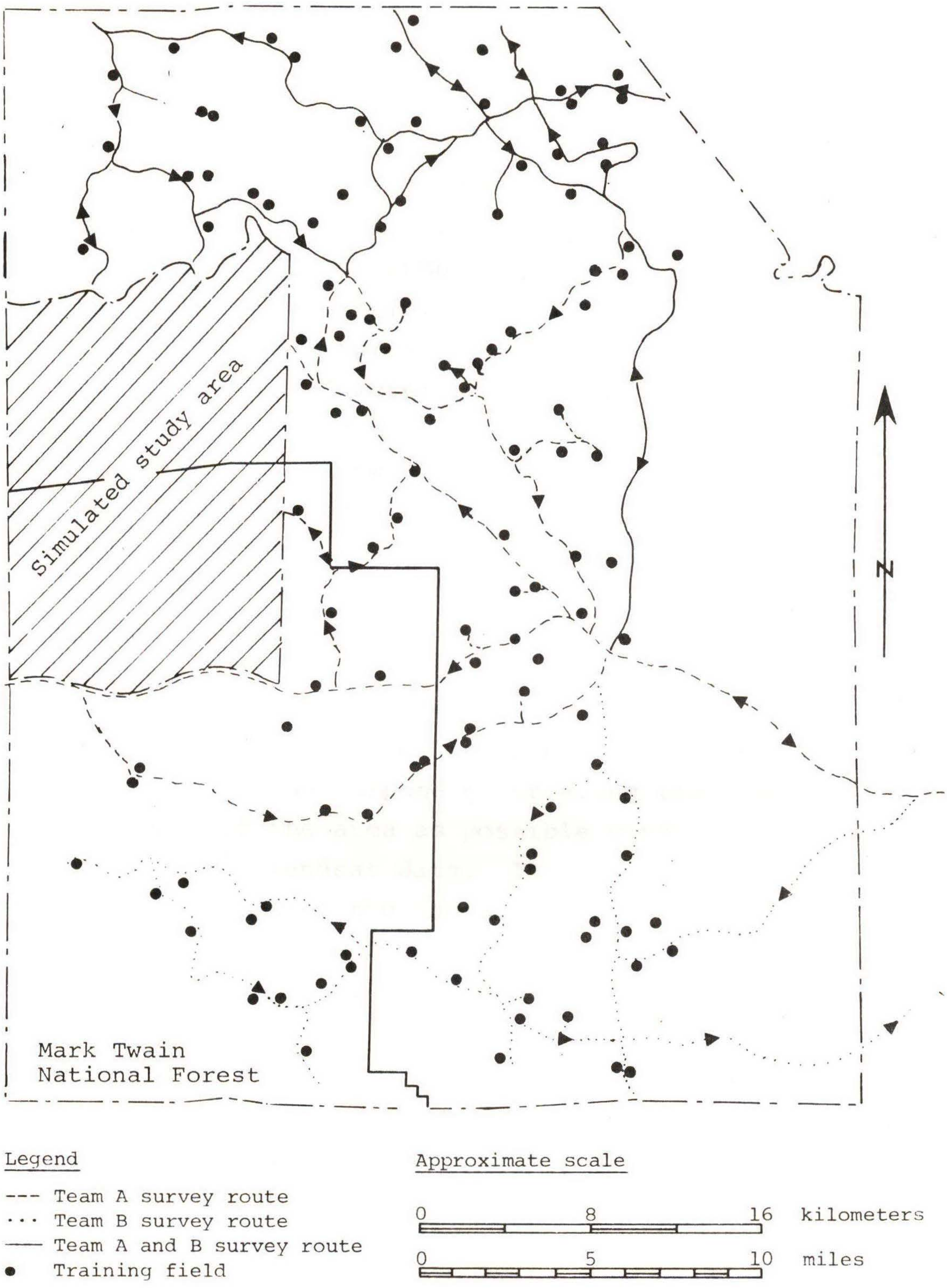


Figure 4-1.— Washington County, showing administrative boundaries, training fields, simulation test area, and survey routes.

5. PREPROCESSING

5.1 OVERVIEW

The preprocessing task involved the two Landsat data sets selected as explained in appendix A; i.e., July and November 1974. A third data set was generated by combining Landsat radiometric bands 5 and 7 from the November acquisition with the same bands from the July acquisition. An area approximately 1000 pixels by 1000 pixels or approximately 56 kilometers by 79 kilometers (35 miles by 49 miles) was delineated on each data set. The temporal data set was compiled by making an image-to-image registration using the ERIPS. The three data sets were then registered to ground control points using a least squares fit program executed on the Programmed Data Processor, Model 11/45 (PDP 11/45) computer.

Before processing the data sets, each set was divided into four segments, each 485 pixels by 485 pixels, which is the optimum number of pixels that can be displayed in a square array on the Image 100 screen. The purpose of creating these segments was to display as much of the area as possible without compression or replication of the Landsat data. The final tasks in this phase were the delineation of the administrative boundaries on each segment and the generation of the false-color composites using the PMIS/DAS and the attached film recorder.

5.2 PROCEDURAL RESULTS

Detailed procedures are outlined in section 4 of the technical analysis procedures (ref. 2). The following discussion briefly touches on significant events, processes, and results. Deviations from reference 2 are noted where appropriate.

5.2.1 IMAGE-TO-IMAGE REGISTRATION

Of the two Landsat scenes selected for the temporal image-to-image registration, the July acquisition was selected as the base image, and the November acquisition was registered to control points located thereon. A total of 84 points were selected on the July scene to execute the polynomial fit on the ERIPS. Of the 84 points selected, 31 were rejected because of excessive errors. After the sixth iteration, the errors stabilized with a root mean square error of 0.8 pixel and a maximum error of 1.4 pixels. The registration was accomplished using the autocorrelation option[2] wherein the operator manually selects the first four control points while the remaining points are chosen by the algorithm on the basis of the relative contrast between the two scenes. The registration was later visually checked on the Image 100 by superimposing the November data on the July imagery.

5.2.2 IMAGE-TO-GROUND REGISTRATION

Registration of the Landsat imagery to ground control points was accomplished using a first-degree polynomial of the form

$$u = ax + by + c$$

$$v = dx + ey + f$$

where x and y are the original coordinates, u and v are the new coordinates, c and f are constants representing a shift of origin, and a, b, d, and e are scaling factors. The initial values of the polynomial coefficients were obtained by performing a least squares fit of the coordinates of a set of measured map control points and their corresponding Landsat coordinates. The initial coefficients were then used to derive corrected (best fit)

---

[2]Autocorrelation Sequential Similarity Detection Algorithm (SSDA).

Landsat coordinates. The root mean square error was held to 1 pixel or less; and, if this was not achieved, points showing excessive error were discarded and the adjustment run again. The results of running the registration on the Image 100 provided a rotation factor (shift due to the Earth's rotation), which was used to deskew the Landsat image and display the scene on the Image 100 screen in proper perspective. The selection of the control points was based on high contrast features, such as road intersections or dam sites, which could easily be located on the base map as well as on the Landsat imagery. Mensuration of the map points was accomplished using the Dell Foster coordinatograph, whereas the Image 100 was used to measure the Landsat coordinates of each corresponding identifiable point. Again, the July Landsat acquisition was used as the base because its broad spectrum of hues and tones displayed more detail and hence more identifiable control points.

The results of running the least squares adjustment through five iterations on the PDP 11/45 computer provided a good fit with line and sample root mean square errors of 0.757 pixel and 0.104 pixel, respectively. As a result of applying the registration parameters, the image as displayed on the interactive computer screen was foreshortened from 485 scan lines to only 349 lines. As prescribed procedures called for an aspect ratio of 1:1, this deficiency was corrected by a procedure that replicated approximately every third pixel, thus creating an image of 485 pixels by 485 pixels.

5.2.3 ADMINISTRATIVE BOUNDARIES AND TRAINING FIELDS

The only administrative boundary of any significance other than the county boundary was that of the Mark Twain National Forest, which occupies approximately 30 percent of the county area. The county map shown in figure 4-1 indicates this area as well

as the area selected for the simulated inventory study. The ground-verified training fields and routes are shown to indicate their distribution and proximity to the simulation study area. Because of lack of time or redundant coverage, only 100 of the 122 training fields shown were visited in the field.

6. PROCESSING

6.1 OVERVIEW

The processing procedures outlined in reference 2 required a type separability study for each of the Landsat data sets and also a simulated inventory study using the most accurate data set as determined from the separability study.

For the separability study, training field signatures were developed from the ground-verified training fields and used to determine the separability (consistency) of the radiometric values for each of the hierarchical categories; i.e., softwood, hardwood, grassland, and water. If the accuracies of this Level II landcover classification equaled or exceeded the requirements, shown in table 6-1, a Level III separability study was to be attempted, with the intent of separating the forest species. The results of these studies were then to be used as guidelines for processing the Landsat data for the simulation study.

The final part of the task was directed toward processing the data for the simulated inventory. This study was designed to use training fields selected strictly from aerial photographs with no aid from ground sources; yet it would employ the expertise gained from the previous interpretation done for the separability study, which did have the advantage of comparison with ground data.

6.2 TYPE SEPARABILITY STUDY

6.2.1 TRAINING FIELD SELECTION

Of the 100 training fields verified on the ground, 83 were used in the separability study; the remainder were discarded because

of positional errors or failure to be located on Landsat imagery. The verified fields were categorized into four broad classes:

Softwood — Predominantly shortleaf pine (*Pinus echinata*) with scattered eastern red cedar (*Juniperus virginiana*)

Hardwood — Predominantly oak-hickory and mixed oak consisting of white oak (*Quercus alba*), post oak (*Quercus stellata*), blackjack oak (*Quercus marilandica*), black oak (*Quercus velutina*), northern red oak (*Quercus ruba*), and scarlet oak (*Quercus coccinea*); minor stands of American sycamore (*Platanus occidentalis*), willow (*Salix negra*), and maple (*Acer rubrum*)

Grassland — Predominantly well-developed and improved pastures usually of small area;[3] also includes unimproved grasses and fallow pastures

Water — Includes all water bodies, with no attempt to separate turbid water from clear water spectrally

At the time of the field survey, the dominant and subdominant forest stands, the understory, the soils, the water regime, and the aspect and slope were observed and recorded. The intent of recording these data was to attempt a correlation with the Landsat radiometric values in the hope of detecting variations that might help resolve signature anomalies or provide a means for developing a Level III hierarchy. Therefore, initial efforts were directed toward separating the various admixtures of softwoods

[3]Interpretation of the July Landsat scene indicates most pastures were cut immediately before overflight.

and hardwoods; and, if time permitted, efforts would be made to separate grasslands and water into their respective units (table 6-2). On the basis of the ground data for each training field, the forest admixtures were categorized into dominant and subdominant vegetation, with the mixed oaks grading into admixtures of hickory, pine, and cedar in varying amounts. Table 6-2 shows the Level II hierarchy and the potential Level III hierarchy.

6.2.2 TRAINING FIELD LOCATION

The procedures outlined in reference 2, which required digitizing the training field coordinates and then using the Irregular Cursor program (an Image 100 program), were modified in order to maintain positional integrity[4] and expedite the overall process of locating training fields on the Landsat data. In the modified procedure, the training fields were manually transferred from the aerial photographs to the PMIS/DAS transparencies, and then these transparencies were used as visual aids in relocating the training fields on the Image 100 screen.

6.2.3 SIGNATURE ACQUISITION

Once the locations of the training fields were determined, it was an easy task to drive the Image 100 cursor to its correct position. (The position of the cursor was verified by visual inspection as well as by correlation of the radiometric values with those of other training fields of the same category.) The signatures were acquired using the One-Dimensional Signature Acquisition program and driving the cursor to the location plotted on the PMIS/DAS transparencies. The signature acquisition

---

[4]Previous work conducted at the other TES sites indicated that the manual method provided more accurate positioning and was less time consuming.

program was then executed, the Image 100 display alarmed,[5] and the signature displayed on the Tektronix graphics terminal (fig. 6-1).

As the July data set was the reference for the image-to-image registration process, it was also used as a base for the location of the 83 training fields identified in Washington County. Training field signatures for the November data set and the temporal data set were obtained by overlaying each data set on the July set, moving the cursor to the previously determined position, and executing the signature acquisition program. This technique expedited the procedure and provided consistency of training field positioning from one data set to another.

Because the area had been divided into quadrants, it was necessary first to develop from the training fields in each quadrant a composite signature for each class and then to combine each set of four composite signatures into a "super-signature" for the class. This supersignature was then stored in the Image 100 computer to be used during the separability study.

6.2.4 ASSESSMENT OF TRAINING FIELD SELECTION FOR LEVEL II

As part of the established procedures, an assessment of the selection of the training fields was obtained by performing a maximum likelihood classification using the relative radiance values from the supersignature of each of the classes in the Level II hierarchy. From the results of this classification, it was possible to assess the signature of each training field to determine how representative it was. This determination

---

[5] Term used to describe the classification of a theme derived from a training field while operating the Image 100.

was made by noting the number of correctly classified pixels versus the total number of pixels in each original training field. The accuracies, expressed as the percentage of correctly classified pixels, are shown in table 6-3. The results indicated sufficient accuracy to warrant a Level III separability study.

Before a Level III study was begun, the signatures for the training fields representing the Level III categories were visually compared and found to be without sufficient spectral separation to enable a Level III separability. However, it did appear possible that an intermediate separation could be effected by introducing a mixed category of softwood and hardwood (with pine varying from 50 to 80 percent) while still retaining the other Level II categories. Unfortunately, the results of this attempt reduced the accuracies below the prescribed requirements (table 6-4). Therefore, because of the following factors, it was considered an unnecessary expenditure of processing time to pursue this type of analysis:

1. Landsat signatures from the training fields did not show sufficient spectral separation of the hardwoods to make possible a species or physiognomic separation.

2. Ground observations did not reveal a physical separation that would enable a Level III separation.

3. Admixtures of various species of hardwoods and softwoods could not be correlated with aerial photographs to verify a Level III separation.

6.2.5 ANALYSIS OF SEPARABILITY RESULTS

As can be seen in table 6-3, the accuracy figures for each of the data sets were well within the limits established by reference 2 (table 6-1). These results indicate that the training fields were selected and identified with a great deal

of consistency. The introduction of the mixed class (as shown in table 6-4) lowered the overall accuracy and indicated an inconsistency in the selection of training fields. Of the three data sets, the November acquisition showed the lowest accuracies; the low Sun angle affected the insolation of the training fields in this data set. The temporal data set would of course be influenced somewhat by the same shadows; consequently, the relatively higher values are unexplained.

6.3 SIMULATED INVENTORY STUDY

6.3.1 TRAINING FIELD SELECTION

The simulated inventory study was designed to investigate the accuracy to which the site could be classified using a limited area (10 percent of the site) and only aerial photographs for training field selection. The July data set was used for this study because of the relative absence from this data set of terrain shadows, which were prominent in the November data set, and because of its much better separability results.

As the inventory was to be made at Level II, no difficulty was encountered in selecting representative training fields for each of the categories. Once the training fields were selected, the processing followed the procedures previously used for the separability study. The signatures from the training fields were combined to form a single composite signature for each of the classes; i.e., softwood, hardwood, and grassland. Unfortunately, the streams in the simulation test area were too small to be delineated as a training field; hence, water was not used as a class.

6.3.2 ASSESSMENT OF TRAINING FIELD SELECTION

Procedures for testing the composite signatures were similar to those used during the separability study, but in this case the signatures were tested against simulated ground truth; i.e., that which was interpreted from aerial photographs of the 10-percent area. The results of this accuracy assessment are tabulated in table 6-5.

6.3.3 INVENTORY RESULTS FOR LEVEL II

The results of the simulated inventory are tabulated in table 6-6. The figures shown represent class proportions based on pixel count.

6.3.4 ANALYSIS OF INVENTORY RESULTS

The results of the assessment of the selection of those training fields used for the simulated inventory study showed an overall accuracy of 99 percent (table 6-5). The increased accuracy for the simulated inventory study over that for the separability study is remarkable considering that only high-altitude photography was used to select the training fields while the separability training fields were selected with the advantage of ground observations. Actually, these accuracies are a measure of the consistency with which the training fields were identified and the signatures acquired for each class. Therefore, these accuracies should not be used to assess the overall classification and inventory. Refer to appendix B for a comparison of the results.

TABLE 6-1.— TRAINING FIELD ACCURACY REQUIREMENTS

Class	Accuracy, percent
Softwood	90
Hardwood	90
Grassland	80
Water	90

TABLE 6-2.— SEPARABILITY HIERARCHY

Level I	Level II	Potential Level III
Forest	Softwood	Pine Cedar
	Hardwood	Upland: Mixed oak Mixed oak and hickory Mixed oak, hickory, and pine-cedar Lowland: Mixed sycamore, poplar, maple, and willow
Nonforest	Grassland	Improved Unimproved
	Water	Clear Turbid

TABLE 6-3.– TRAINING FIELD ACCURACIES

Class	Separability study for Level II		
	July, percent	November, percent	Temporal, percent
Softwood	94	87	94
Hardwood	100	97	100
Grassland	100	98	100
Water	100	84	100
Overall accuracy	98	92	98

TABLE 6-4.– TRAINING FIELD ACCURACIES
WITH MIXED CATEGORY

Class	Separability study for Level II		
	July, percent	November, percent	Temporal, percent
Softwood	77	73	88
Hardwood	77	72	67
Mixed softwood-hardwood	54	49	68
Grassland	99	96	99
Water	78	100	100
Overall accuracy	78	73	79

TABLE 6-5.— SIMULATION TRAINING FIELD ACCURACIES

Class	Accuracy, percent
Softwood	98
Hardwood	100
Grassland	100
Water	Not evaluated

TABLE 6-6.— SIMULATED INVENTORY RESULTS

[Class proportions based on pixel count]

Class	Proportion
Softwood	0.0681
Hardwood	.4451
Grassland	.0980
Water	Not evaluated
Other	.3888

*** ONE DIMENSIONAL TRAINING -REV. A- ***

CHANNEL	EFFECTIVE RESOLUTION	BOUND-THRESHOLDS(X) LOW	INC/D	HIGH	START LEVEL	FINISH LEVEL
1	64	0.00	(100.00)	0.00	0	63
2	64	0.00	(100.00)	0.00	0	63
3	64	0.00	(100.00)	0.00	0	63
4	32	0.00	(100.00)	0.00	0	31

/// TRAINING IN PROGRESS
TRAINING COMPLETE ///

CHANNEL	SPECTRAL-BOUNDS	DELTA	PEAK	MEAN	VARIANCE
1	(11-15)	5	8.	13.00	0.83
2	(6-10)	5	5.	6.92	1.24
3	(18-25)	8	5.	21.83	2.97
4	(11-15)	5	4.	12.67	1.56

TRAINING AREA= 12.
ALARMED AREA= 81640. (31.1%)
PARALLELEPIPED CELLS= 1000.
FIGURE OF MERIT= 8389.

Figure 6-1.– Tektronix display showing results of one-dimensional training. (The mean signature is that of softwood.)

7. POSTPROCESSING

7.1 OUTPUT MATERIALS

The postprocessing activity consisted of outputting materials to be used for the evaluation phase and the generation of a final map product. Specifically, the following materials were produced:

1. One magnetic tape containing four files (one for each quadrant) of the simulated inventory for July 1974
2. One alphanumeric classification map of the simulated inventory for July 1974 (produced on the Gould printer/plotter)
3. Two sets of PMIS/DAS transparencies showing the simulated inventory classification results

The Gould alphanumeric classification map and the PMIS/DAS transparencies were needed for the evaluation of the simulated inventory study. The computer classification tape containing the final classification data of the simulated inventory study was prepared for later shipment to Seiscom Delta, which had contracted to produce a color-coded classification map at a scale of approximately 1:350 000. This map was to be produced by a direct image process from the computer tape to a light-sensitive plate via a laser beam recorder.

7.2 PRIMARY SAMPLING UNIT LOCATION

Each primary sampling unit (PSU) to be used in the evaluation phase consisted of a 50-pixel by 50-pixel area. Within each PSU, 10 secondary sampling units (SSU's) were randomly selected, each consisting of an area 2 pixels by 2 pixels in size. The PMIS/DAS transparencies were then used to assist in precisely locating the PSU's. This was accomplished by randomly selecting the PSU's on the aerial photographs, transferring these areas to the PMIS/DAS transparencies by visual inspection, and finally

transferring them to the Gould output classification map. This last step was accomplished by using the Dell Foster coordinatograph to obtain the corner coordinates of each PSU and then applying a simple coordinate transformation to derive the corner coordinates of the PSU's for the Gould classification map. These coordinates were then plotted manually.[6]

[6]This process could have been accomplished in a more efficient manner by simply locating the PSU on the Image 100 screen before generating the Gould classification map.

8. EVALUATION

The evaluation phase of the processing was designed to measure the accuracy of the simulated ADP classification, which used training fields that had been selected from a relatively small area (10 percent of the site) and which used only photointerpretation skills. The objectives were (1) to test the accuracy of the ADP classification by performing a statistical analysis, deriving class proportion estimates from limited random sampling and calculating a PCC; (2) to compare the class proportion estimates to a credible inventory baseline; and (3) to perform a linear regression analysis in order to improve the estimates of proportions and to reduce the variance of the estimates. (For this evaluation, the July data set was used because it showed a relatively high correlation with ground truth and because, unlike the November and temporal data sets, it was unaffected by the low Sun angle and resulting shadows.)

8.1 STATISTICAL ANALYSIS

8.1.1 PROCEDURES

The first objective was achieved by the random selection of PSU's and SSU's on the Gould classification map. This map was then correlated with the aerial photographs. As each PSU was alined with the photograph, the areas covered by the SSU's underwent photointerpretation, giving class proportions for the PSU. Class proportion estimates for the PSU were also made by tabulating the class symbols in the SSU's on the Gould classification map. At the same time, the Gould classification of each SSU was compared to the photointerpreted classification to determine its correctness. A PCC_i ($i = 1, 2, \cdots, 10$) was calculated for each PSU by dividing the number of correctly classified SSU's by the 10 SSU's in a PSU. These PCC_i's were then averaged to produce

an overall map accuracy PCC. (For more details pertaining to the statistical evaluation, see reference 2.)

8.1.2 RESULTS

As a result of the statistical evaluation, a relative classification accuracy was obtained – relative to the interpretation of the aerial photography. This evaluation did not consider the positional accuracy of the classification; nevertheless, it did attempt to account for the residual registration error by taking the best fit of a series of nine possible locations (see ref. 2) surrounding the calculated position of each SSU. The best fit is defined here as the location of the SSU that best corresponds with the photointerpretation of the ground truth for that SSU.

According to reference 2, the evaluation procedure considers the extent of feature classes on the map; that is, the PCC is weighted by the relative occurrences of the classes and is a measure of the overall accuracy of the map rather than of accuracies of individual feature classes.

Table 8-1 gives a summary of the PCC calculations for 10 PSU's. For each PSU, a PCC_i ($i = 1, 2, \cdots, 10$ was obtained by comparing the classification of each SSU against the corresponding area on the aerial photograph, as explained above. The mean PCC was calculated to be 85 percent. To this was applied the half confidence interval of 6.2 percent (computed at a level of 0.90), giving a confidence spread of 78.8 to 91.2 percent. Table 8-2 presents a summary of class proportions for each of the PSU's, where p_i represents the proportion based on photointerpretation and $\hat{p}_i$ that proportion obtained from the simulated inventory classification. For each category, average proportions were obtained for both p_i and $\hat{p}_i$ and were used in calculating the proportion errors shown in table 8-3.

8.1.3 DISCUSSION

8.1.3.1 Softwood

Table 8-2 shows the softwood column to be deficient in sampling data as compared to the adjacent column for hardwood. This lack of softwood samples is undoubtedly due to the distribution of the softwood stands; they are generally well dispersed, with a weak trend toward increased density in the west-central portion of the site.

With softwood occupying such a minor portion of the area, the sampling was not dense enough to encompass such widely dispersed stands. Then, too, the softwood stands are not uniform but grade from small and isolated stands through moderately dense stands interspersed with hardwood to very dense plantation stands. On the basis of this distribution, it would appear that a systematic method of sampling would gain more information per PSU than a purely random method.

8.1.3.2 Hardwood

The profusion of hardwood provided adequate sampling (table 8-2), showing a fair correlation of the photograph sample proportion with the inventory sample proportion. The inventory estimate for PSU 5 is the only one deviating any appreciable amount from the photograph sample estimate. It is believed that the overall classification could have been improved if larger training fields had been selected. Although this would have increased the variance of the hardwood signature, it could have captured more hardwood stands without seriously saturating the class with other categories whose signatures may overlap the expanded signature.

8.1.3.3 Grassland

For grassland, the correlation of the photograph samples with the inventory samples (table 8-2) was not satisfactory. This

discrepancy was not anticipated, as usually little difficulty is encountered when classifying a high-contrast area such as grassland. The overall statistics shown in table 8-3 for this class are marginal, as indicated by the large relative error of 14.75 percent. This sizable error is no doubt due to the great variance in the signatures of the various types of grasses and the fact that many of the fields had been mowed, giving rise to a greater variety of signatures. Obviously, many of these samples found their way into the "other" category if they were not within the bounds of a grassland training field signature. Probably the major factor contributing to the large error was the limited amount of grassland in the 10-percent area from which the representative training fields were selected. Only a very small percentage of the test area was grassland.

8.1.3.4 "Other"

Although the category "other" is not a class as such, it requires consideration because many of the class deficiencies seem to have occurred because of migration toward this catchall category. Figure 8-1 helps to illustrate this effect. When the inventory proportions are compared to the photointerpretation proportions, it is interesting to note that each of the major classes (i.e., softwood, hardwood, and grassland) has forfeited a portion of its contents — none has gained additional pixels.

8.1.4 ANALYSIS OF RESULTS

According to the procedures designed for TES, the simulated inventory achieved acceptable limits by providing a PCC of 85, with a spread from 78.8 percent to 91.2 percent at the 90-percent confidence level. However, the excess of pixels under the category "other" requires further analysis. Several factors could contribute to this anomaly: (1) the signature of the

"other" category is extremely broad as it contains all other categories that were not in the classification hierarchy (i.e., water, urban, roads, bare soil, and shadows), and thus it could ensnare any unidentified pixels from the three classes; (2) the effects of apparent shadows could confuse the classification of a northwest-facing slope and place it in the "other" category (it is not known just how much the combination of Sun angle and hilly terrain affects the signatures involved, but it is believed to be worthy of further investigation).

8.2 COMPARISON OF INVENTORY RESULTS TO COUNTY STATISTICS

The second objective of the evaluation was achieved by comparing the simulated inventory results to forest statistics for the county obtained by the USDA Forest Service from a very accurate survey conducted in 1972. The results of this comparison are tabulated in table 8-4 (inventory class proportions are based on a pixel count).[7]

When the two sets of data are compared, the softwood estimate shows an excellent correlation but the hardwood estimate is extremely deficient. Here again, the insufficiency of training field signatures to represent a cross section of all the admixtures of hardwoods resulted in a migration of pixels to the "other" category.

8.3 REGRESSION ANALYSIS

In order to reduce the variance of the estimates, a linear regression transformation was determined using the sampled class proportion ($\hat{p}$) versus the photosampled class proportion (p) of

[7]During the processing phase, additional data were obtained to provide a relative means of comparing the simulated inventory class proportion estimates to the estimates obtained from the separability study. See appendix B for the results of this comparison.

the simulated inventory. The regression analysis was applied to the class proportions taken from the entire simulated inventory ($\hat{p}_{inv}$), and the resulting estimates are shown in table 8-5. The results of the regression indicate significant bias in the proportion estimates of the ADP classification compared to the photointerpreted results. See appendix C for further details on the regression analysis.

TABLE 8-1.— SUMMARY OF PCC CALCULATIONS

PSU number	PCC_i expressed as proportion
1	0.80
2	.90
3	.80
4	.90
5	.70
6	.90
7	1.00
8	.80
9	1.00
10	.70

$$PCC = \frac{1}{m} \sum_{i=1}^{m} PCC_i$$

$$= 0.85$$

$$S^2_{PCC} = (1 - f) \frac{1}{m(m - 1)} \sum_{i=1}^{m} (PCC_i - PCC)^2$$

$$S_{PCC} = 0.034$$

$$\Delta = tS_{PCC}$$

$$= 1.833S_{PCC} \text{ at 0.9 confidence interval}$$

$$= 0.062$$

$$\text{Confidence interval of PCC} = (PCC - \Delta,\ PCC + \Delta)$$

$$= (0.788,\ 0.912)$$

Inventory PSU's	PCC	Half confidence interval at 0.9	PCC ± $\Delta_{0.9}$
10	85%	±6.2%	(78.8% - 91.2%)

Notation

m = number of PSU's in sample scheme

PCC_i = percent correct classification (i = PSU index)

f = finite population constant = (m - 1)/(N - 1), where m = number of PSU's in sample scheme and N = total number of PSU's in entire population

S^2_{PCC} = variance of mean

S_{PCC} = standard deviation

t = constant obtained from statistical tables

TABLE 8-2.– SUMMARY OF CLASS PROPORTIONS

PSU number	Softwood		Hardwood		Grassland		Water[a]		Other	
	p_i	$\hat{p}_i$	p_i	$\hat{p}_i$	p_i	$\hat{p}_i$	p_i	$\hat{p}_i$	p_i	$\hat{p}_i$
1	0.00	0.000	0.85	0.750	0.14	0.125			0.01	0.125
2	.00	.025	1.00	.825	.00	.000			.00	.150
3	.00	.000	.77	.875	.13	.000			.10	.125
4	.00	.050	.96	.850	.02	.000			.02	.100
5	.09	.100	.81	.625	.08	.000			.02	.275
6	.07	.050	.82	.775	.10	.100			.01	.075
7	.08	.025	.86	.875	.00	.000			.06	.100
8	.00	.000	.95	.950	.00	.025			.05	.025
9	.00	.000	.88	.775	.10	.200			.02	.025
10	.05	.000	.85	.725	.04	.075			.06	.200
Total	0.29	0.250	8.75	8.025	0.61	0.525			0.35	1.200

Average proportion	Softwood		Hardwood		Grassland		Water		Other	
	p	$\hat{p}$	p	$\hat{p}$	p	$\hat{p}$	p	$\hat{p}$	p	$\hat{p}$
	0.029	0.0250	0.875	0.8025	0.061	0.0525			0.035	0.120

[a]None in test area.

Notation

p_i = photograph sample proportion for *ith* PSU

$\hat{p}_i$ = inventory sample proportion for *ith* PSU

p = average photograph sample proportion

$\hat{p}$ = average inventory sample proportion

TABLE 8-3.— SUMMARY OF CLASS PROPORTION ERRORS

Class	Inventory class proportion, $\hat{p}$	Photograph class proportion, p	Average error, B	Standard deviation of error, S_B	Half confidence interval, $\Delta_{0.9}$	Confidence interval, $B \pm \Delta$	Percent relative error, RB
Softwood	0.0250	0.029	0.004	0.0099	0.018	(-0.114, 0.022)	13.79
Hardwood	.8025	.875	.073	.028	.052	(.021, .125)	8.34
Grassland	.0525	.061	.009	.019	.036	(-.027, .045)	14.75
Water[a]							
Other	.120	.035	-.085	.026	.047	(-.132, .038)	-242.86

[a]None in test area.

Notation

$B_i = p_i - \hat{p}_i$ = individual error

$B = \frac{1}{m} \sum_{i=1}^{m} B_i$ = average error

$S_B^2 = \frac{(1 - f)}{m(m - 1)} \sum_{i=1}^{m} (B_i - B)^2$ = variance

$\Delta_{0.9} = 1.833 S_B$ = half confidence interval

$RB = \frac{B}{p} \times 100$ = relative error

TABLE 8-4.— SIMULATED INVENTORY VERSUS WASHINGTON COUNTY STATISTICS

[Class proportions based on pixel count]

Data source	Level II class				Total area, including water
	Softwood	Hardwood	Grassland	Other	
Simulated inventory:					
Proportion	0.0681	0.4451	0.0980	0.3380	1.00
Square hectometers	13 606	88 983	19 587	77 720	199 858
Acres	33 620	219 800	48 400	192 050	493 860
County statistics:[a]					
Proportion	0.0622	0.6743	0.1399	0.1236	1.00
Square hectometers	12 262	133 020	27 596	23 957	197 260
Acres	30 300	328 700	68 190	59 200	487 440
Proportion error	0.0059	-0.2292	-0.0419	0.2652	0.013

[a]County statistics obtained from references 4, 8, and 9.

TABLE 8-5.— REGRESSION ESTIMATES OF PROPORTIONS AND ASSOCIATED PRECISION

Class	Simulated inventory proportion, $\hat{p}_{inv}$	Regression equation	Regression estimate of proportion, $p\|\hat{p}_{inv}$	Coefficient of determination, r^2	Variance of the regression estimate, s^2	Half confidence interval, $\Delta_{0.9}$	Percent relative variance, $\frac{\Delta_{0.9}}{p\|\hat{p}_{inv}} \times 100$
Softwood	0.0681	$p = 0.725\hat{p}_{inv} + 0.011$	0.060	0.390	0.000261	0.030	49.4
Hardwood	0.4451	$p = 0.330\hat{p}_{inv} + 0.610$	0.757	0.174	0.007809	0.162	21.4
Grassland	0.0980	$p = 0.417\hat{p}_{inv} + 0.039$	0.080	0.278	0.000326	0.030	41.4

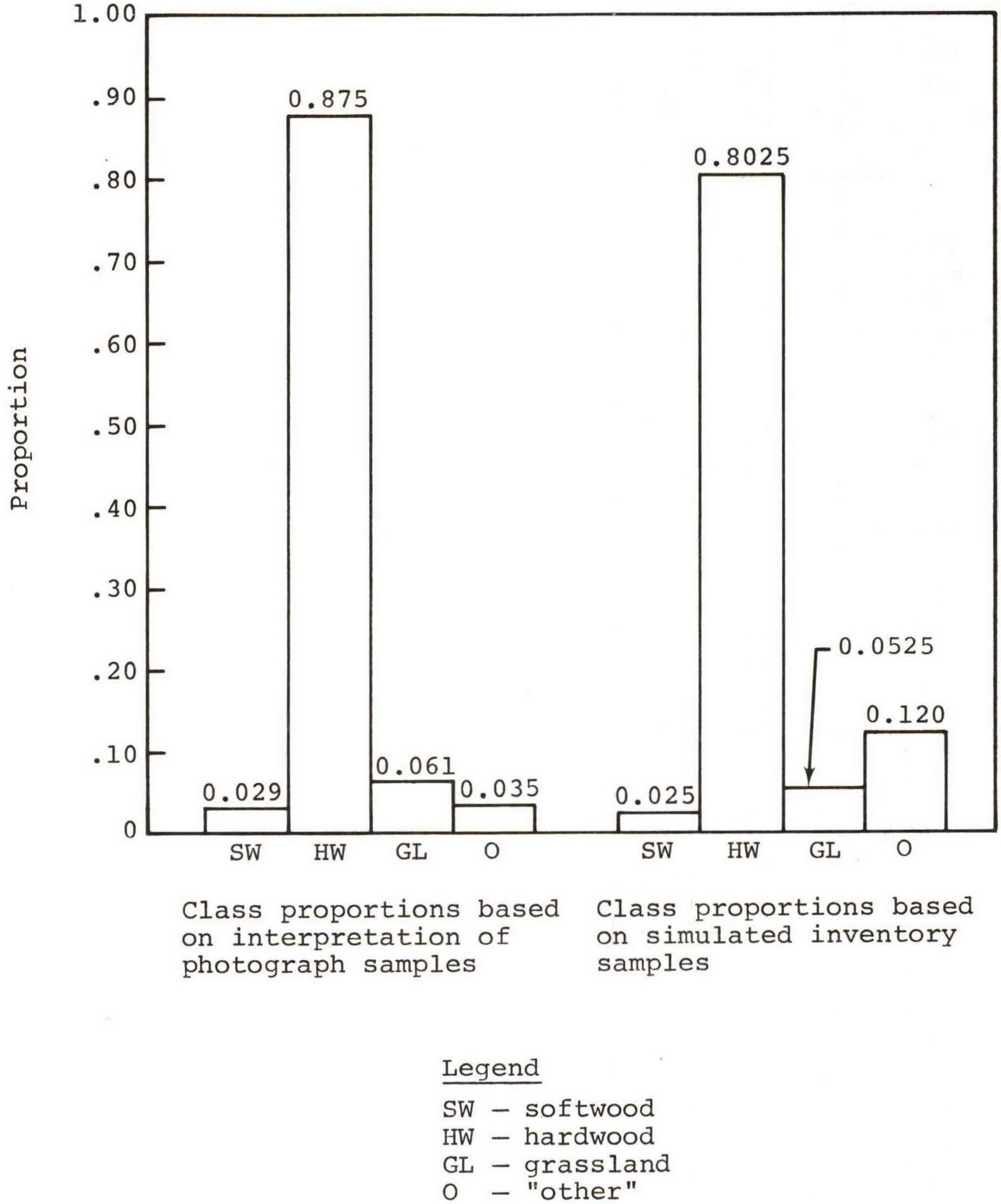


Figure 8-1.– Graphic display of class proportion estimates.

9. RESOURCE UTILIZATION

Table 9-1 is a tabulation of the various tasks and actual machine hours associated with the analysis and processing of the Washington County TES site. Table 9-2 shows the equipment used and the resulting costs. The costs of transportation to and from the site, food, and lodging for the four people who visited the site are shown as trip expenses.

The direct cost of \$46 947 divided by the total area processed shows a cost of \$0.158 per square hectometer (\$0.064 per acre). Work done in a production mode would presumably be less costly because the analysts would be more familiar with the system procedures. However, it is anticipated that future work such as this may increase in cost as the rate of inflation continues to rise despite a reduction in man-hours and computer usage.

Resources required for the study but not included in the overall costs are items such as the cost of eight Landsat tapes and corresponding photographic color composites and the cost of the CIR aerial photographs from Mission 289 flown by NASA in September and November of 1974.

TABLE 9-1.— RESOURCES UTILIZED FOR SITE PROCESSING

Task	Man-hours	Actual machine hours				
		ERIPS registration	Image 100 interactive analysis	PMIS/DAS image composition	Dell Foster digitizing	Univac product merge and cleanup
Preliminary image analysis	140					
Site analysis (including trip)	303					
Preprocessing	180	6	30	10	2	0
Processing	129		46	8		
Postprocessing	90		12	6	8	0
Evaluation	88					
Reporting	256					
Total	1186	6	88	24	10	0

TABLE 9-2.– DIRECTS COSTS FOR MACHINE HOURS AND MAN-HOURS

Item	Cost per hour	Hours	Total cost
Machine:			
ERIPS	$300.00	6	$ 1 800
Image 100	300.00	88	26 400
PMIS/DAS	100.00	24	2 400
Dell Foster	15.00	10	150
Univac 1110	300.00	0	0
Total machine cost			$30 750
Man-hours	$ 12.43	1186	$14 742
Trip expenses			$ 1 455
Total direct cost			$46 947
Cost per square hectometer[a]			$ 0.158
Cost per acre[a]			$ 0.064

[a]These figures apply to the separability study and are based on the whole area processed, which extended beyond the county boundaries and equaled 297 367 square hectometers (734 794 acres).

10. CONCLUSIONS

The Washington County, Missouri, study was implemented with the intent of achieving the objectives outlined in the Ten-Ecosystem Study (TES) Investigation Plan; namely:

1. To perform a type separability study to determine the possible detail to which forest and grassland can be mapped from satellite data
2. To conduct a simulated inventory study to determine how successfully ADP technology can extend limited ground truth to make large-area inventories
3. To establish the best season for accurately mapping each site or ecosystem

10.1 TYPE SEPARABILITY STUDY

On the basis of the assessment of the accuracy of the selection of the training fields used in the separability study, it was concluded that

1. Sufficient spectral separability existed between softwood, hardwood, and grassland to develop a Level II classification and inventory.
2. Interjecting a mixed category in order to account for all of the forested area and achieve a Level III classification reduced the accuracy of each category and consequently the overall accuracy of the classification.
3. The preliminary photointerpretation conducted in the laboratory before the field survey indicated that the differentiation of hardwood and softwood was greatly facilitated by the use of the September and November aerial photographs. However, neither data set alone nor the two together contained sufficient information to allow the analyst to identify positively species types or even to isolate mixed types with any degree of assurance.

10.2 SIMULATED INVENTORY STUDY

The results from the statistical analysis of the July simulated inventory indicate that, relative to photointerpretation, this inventory achieved an accuracy of 85 percent ± 6.2 percent. However, the results from the regression analysis, including some ground verification, indicate significant bias. On the other hand, the comparison of this inventory with county statistics produced acceptable results. These differences in evaluation results indicate that the sampling strategy was not entirely satisfactory.

In addressing the extendability of limited ground truth to make large-area inventories, pertinent questions arise as to the definition of "large area" and also the means of measuring the degree of success. Suffice it to say that the large area in this study was only 196 843 square hectometers (486 400 acres). In assessing the results of the simulated inventory for Washington County, it was found that only softwood and grassland signatures could be extended across the county with less than a 5-percent error (table 8-4). With minor modifications to the techniques used in selecting the training fields (increasing the number and the size), the hardwood signatures could no doubt reach the same degree of accuracy.

10.3 BEST SEASON FOR LANDSAT ACQUISITION

These studies indicate that the best seasons for conducting this type of forest inventory are the fall and summer. If only one date was to be used and a separation of only hardwood and softwood was desired, then the fall season would provide the most separation and the best ADP classification. The optimum time would be immediately after leaf drop while the Sun remains at the highest possible angle.

11. RECOMMENDATIONS

The preceding results and discussions support the following recommendations:

1. The sampling strategy should be modified to increase overall accuracy; a sample density of 0.09 percent is insufficient for TES sites with sparse or widely dispersed forest stands.
2. The influence of shadows, clouds, apparent shadows, and terrain slopes should be considered and compensated for in the experimental design.
3. New techniques should be included in the experimental design; among them might be ratioing of the Landsat spectral bands to increase the separation between classes and perhaps facilitate the separation of forest species.
4. Means should be devised to reduce the replication of pixels before classification.
5. The image-to-ground registration should be modified so that the ground control points (map or photograph points) can be displayed on the Image 100 screen and compared to their true position.

In a review of the results of the various inventories and proportion estimates, one factor that appears to be lacking is an accurate and reliable ground-truth data base. Such a data base would encompass sufficient area to derive dependable conclusions. For example, the PSU's could have been evaluated in toto instead of using only a very small percentage of the area; i.e., 10 SSU's totaling 40 pixels of the 2500 available in each PSU. This method would not have increased the cost to any appreciable extent.

APPENDIX A

INTERPRETATION AND EVALUATION OF AERIAL PHOTOGRAPHY AND LANDSAT IMAGERY

APPENDIX A

INTERPRETATION AND EVALUATION OF AERIAL PHOTOGRAPHY AND LANDSAT IMAGERY

A.1 AERIAL PHOTOGRAPHY INTERPRETATION

Before the field survey (onsite inspection of forest types) was conducted, the CIR aerial photography (1:120 000 scale) was studied to determine the film signatures of the major classes of vegetation; i.e., softwoods, hardwoods, and grasslands. The image analysts made these determinations using their interpretation abilities coupled with collateral site-specific information.

The identification of the vegetative classes and delineation of their boundaries were based on hue, tone, and texture in addition to the general location of the vegetative class. On the basis of these factors, a preliminary hierarchy was established (table A-1) and used as a guide to delineate on an overlay the boundaries for each category so defined. The September 1974 photography was used for the primary base, as the leaf-on condition would theoretically allow the analyst to delineate boundaries between the upland and lowland hardwoods. Observations during the field survey indicated that this interpretation could not be confirmed. In all probability, the delineated boundaries indicated a physiognomic change rather than a species change. The November photography, showing a leaf-off condition, was used to confirm the existence of softwoods within the hardwood forests and to approximate the density of the softwood stands without the interference of the hardwood canopy. Areas in which the softwoods and hardwoods could not be differentiated were classed as mixed.

After the mapping and interpretation had been completed, training fields to be used for the separability study were identified. The

number of training fields selected to represent each vegetative class was based on the percentage of area covered by each class. To ensure signature integrity, care was exercised to choose training fields that were both representative of the class and relatively homogeneous. Accessibility to the ground inspection crew was a prime criterion in the selection of each training field. Cutover areas, strip mines, and urban areas were identified for later ground inspection and signature analysis to determine possible signature overlap with the prime vegetative classes.

When sufficient training fields had been selected, the area for the simulated inventory was delineated. This area was located in the western portion of the county, with approximately one-half in the Mark Twain National Forest and one-half on private lands. This mixture of private and public lands was chosen to reduce any bias that might occur as a result of different forest management practices and also to ensure future accessibility and forest integrity. The delineated area was approximately 10 percent of the county land area.

A.2 LANDSAT IMAGERY EVALUATION

An interpretive analysis was made of the eight Landsat frames and this analysis was compared to that of the high-altitude photographs. The aerial photographs were considered the control data, and the Landsat imagery was used as the test data.

A.2.1 PROCEDURES

With the aid of the Old Delft stereoscope, a hierarchy of vegetation classes, including features of the Level II category, was selected and identified on the CIR aerial photographs. The corresponding Landsat imagery was interpreted and results recorded. This was accomplished for each of the eight Landsat scenes. The

results were compared and the Landsat scenes ranked according to the highest PCC. The feature classification was based upon tonal, textural, and spatial properties normally used in conventional interpretation exercises. To ensure accurate positioning of photograph features with Landsat features, a control grid system was constructed, and the series of sample points (100) that had been identified on the CIR photographs were then correlated with the Landsat imagery. This correlation was accomplished using the Bausch & Lomb zoom transfer scope and enlarging the Landsat imagery to a maximum.

An overall correlation accuracy (PCC) of softwood, hardwood, grassland, and "other" was calculated from the 100 randomly selected points. The PCC's for hardwood and grassland were calculated but not used because the numbers obtained were very similar and did not provide conclusive results.

A.2.2 RESULTS

The following represents the PCC's from the analysis of the Landsat frames:

Date	Image identification	PCC
May 24, 1973	1305-16121	76.03
August 4, 1973	1377-16111	(8)
July 30, 1974	1737-16025	80.16
September 22, 1974	1791-16004	76.03
October 10, 1974	1809-16002	80.99
November 15, 1974	1845-15591	82.64
March 3, 1975	1953-15544	78.51
May 14, 1975	5025-15511	(8)

[8] Not evaluated (cloud cover).

The July 1974 and November 1974 Landsat images were selected as being the optimum set, indicated by the high correlation with the aerial photography. In addition to the high correlation, there was sufficient temporal separation to indicate a phenological change. It was later determined from field observations that the November date was excellent for discriminating pine from hardwoods because of the total leaf-off condition at the time of the overflight. It should be noted that, although the Landsat scene dated October 10, 1974, exhibited the second highest correlation percentage, it was not selected as there was not sufficient temporal separation.

TABLE A-1.– PRELIMINARY HIERACHY[a] FOR WASHINGTON COUNTY

Class	Estimated percentage of area	Number of training fields[b]
Softwood	4	2
Hardwood:		
Upland hardwood	35	52
Lowland hardwood	26	39
Agriculture-grassland	20	20
Cutover areas	5	2
Water	2	3
Urban (Potosi)	1	1
Mixed hardwood-softwood	5	2
Mining (surface)	2	1
Total	100	122

[a]The hierarchy used in defining potential training fields on the high-altitude CIR photographs.

[b]The number of training fields allocated to each class based on area estimates.

APPENDIX B

COMPARISON OF SIMULATED INVENTORY RESULTS WITH TYPE SEPARABILITY RESULTS

APPENDIX B

COMPARISON OF SIMULATED INVENTORY RESULTS WITH TYPE SEPARABILITY RESULTS

B.1 TRAINING FIELD ACCURACIES

As an added check of the accuracy of the training fields used in the simulated inventory study, their composite signatures were evaluated against those for test fields whose classifications were verified by ground observations. (These test fields were identical to the training fields used in the separability study.) The results were quite different from those shown in table 6-5; they are tabulated in table B-1 as test B with the results of the previous assessment for the simulated inventory study which are labeled test A.

As is indicated in tables 6-3 and 6-5, the analyst was extremely consistent when selecting training fields for the type separability study and for the simulated inventory study. However, when signatures from the simulation training fields were used to alarm the test fields from the separability study, as in test B, it became apparent that the two sets of signatures were not identical. Obviously, the explanation for this discrepancy lies in a more thorough analysis of each set of signatures. The questions that must be explored are these: (1) Do these tests represent an accurate assessment of the selection of the training fields for the simulated inventory? (2) Is the poor correlation of the two tests caused by signature extension over too great an area? (3) If signature extension is the problem, what is the maximum distance the signature can be extended without significant loss of signature integrity? (4) Is the poor correlation between the simulation and separability training fields caused by a physiognomic change or misidentification? Item 4 may prove to hold the most probable explanations for the apparent poor correlation.

B.2 INVENTORY RESULTS

To discover the effects of the different means of verifying training fields on the class proportion estimates made from them, the simulated inventory was expanded to cover the whole 970-pixel by 970-pixel area. (This was the area covered by the separability study.) In addition, the three data sets used in the separability study were used to obtain class proportion estimates. The results are shown in table B-2.

B.3 COMPARISON OF INVENTORY RESULTS

Table B-3 shows the results of the larger simulated inventory and the three separability inventories, along with their training field accuracies, in comparison to the results of the inventories made by random sampling of the simulation classification and by photograph sampling of the same.

In attempting to compare the six types of inventories of table B-3, the method of training field verification must be considered as well as the resultant values and proportion estimates. The following outline attempts to correlate and highlight the subtle differences between the types of inventories:

1. The simulated inventory A shows relatively high training field accuracies comparable to those of separability inventories B and D, yet the class proportions are widely different.

2. For the separability inventories B, C, and D, the class proportions vary considerably even though ground-verified training fields were used.

3. While the training field accuracies of inventory E are satisfactory, the class proportion estimates vary considerably from the values shown in the other inventories. It should be noted that the estimates for inventories E and F were obtained by sampling only a small portion of the total classification.

4. The proportion estimates for inventory F are considered to be more accurate than the values obtained from the computer classifications as they were obtained by interpretation of the aerial photography. However, these values are somewhat biased as the class proportions (derived from sampling) are based on only 0.09 percent of the total pixels classified.

In summary, it can be speculated that the large variations in class proportion estimates are due to a number of factors, the most significant being the admixture of pine with hardwood. The softwood intermingles with the hardwood stands in varying proportions throughout the site. This is not readily apparent from the aerial photographs; hence, the proportion estimates of softwood versus hardwood are difficult, if not impossible, to obtain except where the softwood occurs in nearly pure stands (on plantations) or is dominant enough to influence the CIR photograph. A second factor that contributed to reducing the overall accuracy was the selection of too few training fields in the grassland category, producing a signature that was too definitive and causing much of the grassland to be categorized as "other."

TABLE B-1.— ACCURACY OF SIMULATION TRAINING FIELD SIGNATURES AS MEASURED AGAINST AERIAL PHOTOGRAPHS (TEST A) AND GROUND TRUTH (TEST B)

Class	Test A, percent	Test B, percent
Softwood	98	73
Hardwood	100	82
Grassland	100	77
Water	Not evaluated	Not evaluated
Overall accuracy	99	77

TABLE B-2.— SEPARABILITY AND SIMULATED INVENTORY PROPORTIONS

[Class proportion values based on pixel count]

Class	July simulated inventory	Separability		
		July	November	Temporal
Softwood	0.064	0.055	0.056	0.031
Hardwood	.441	.575	.750	.547
Grassland	.102	.135	.119	.071
Water	None	.001	.002	.001
Other	.393	.234	.072	.350

TABLE B-3.– SUMMARY OF LEVEL II ESTIMATES

Type of inventory	Training field accuracy and class proportion[a]				
	Softwood	Hardwood	Grassland	Water	Other
A. July 1974 simulated inventory	98% / 0.064	100% / 0.441	100% / 0.102	(b) / (b)	(b) / 0.393
B. July 1974 separability inventory	94% / 0.055	100% / 0.575	100% / 0.135	100% / 0.001	(b) / 0.234
C. November 1974 separability inventory	87% / 0.056	97% / 0.750	98% / 0.119	84% / 0.002	(b) / 0.072
D. Temporal separability inventory	94% / 0.031	100% / 0.547	100% / 0.071	98% / 0.001	(b) / 0.350
E. July 1974 inventory by random sampling ($\hat{p}$) of simulation classification	86% / 0.025	92% / 0.802	86% / 0.053	(b) / (b)	(b) / 0.120
F. September 1974 inventory by photograph sampling (p) of simulation classification	(b) / 0.029	(b) / 0.875	(b) / 0.061	(b) / (b)	(b) / 0.035

[a] Accuracy / Proportion

[b] Not applicable.

APPENDIX C

SUMMARY OF REGRESSION ESTIMATES

APPENDIX C

SUMMARY OF REGRESSION ESTIMATES

Figures C-1 through C-3 present graphic plots of the class proportions calculated using different procedures throughout TES. Softwood and hardwood proportions, as well as grassland proportions, are shown for the separability studies. The data are presented in ascending order and by line item; each line is explained in the lower left corner of the figure. It was expected that the separability proportions for the July data would provide the best estimate of the class proportions because the training fields had been selected from throughout the site and verified and were shadow free. However, an examination of the regression estimates in table C-1 indicates a relatively strong correlation between the regression estimates and the November separability proportions.

TABLE C-1.– PROPORTION ESTIMATES

R^2 R^1

Type of inventory	Level II		Level I			
	Softwood	Hardwood	Forest[a]	Grassland	Water	Other[b]
July 1974 simulation $(\hat{p}_{inv})$	0.064	0.441	0.505	0.102		0.393
July 1974 random samples $(\hat{p})$	.025	.802	.827	.525		.120
September 1974 photograph samples (p)	.029	.875	.904	.061		.035
Regression $(p\vert\hat{p}_{inv})$	.057	.756	.813	.082		
July 1974 separability	.055	.575	.630	.135	.001	.234
November 1974 separability	.056	.750	.806	.119	.002	.072
Temporal separability	.031	.547	.578	.071	.001	.350

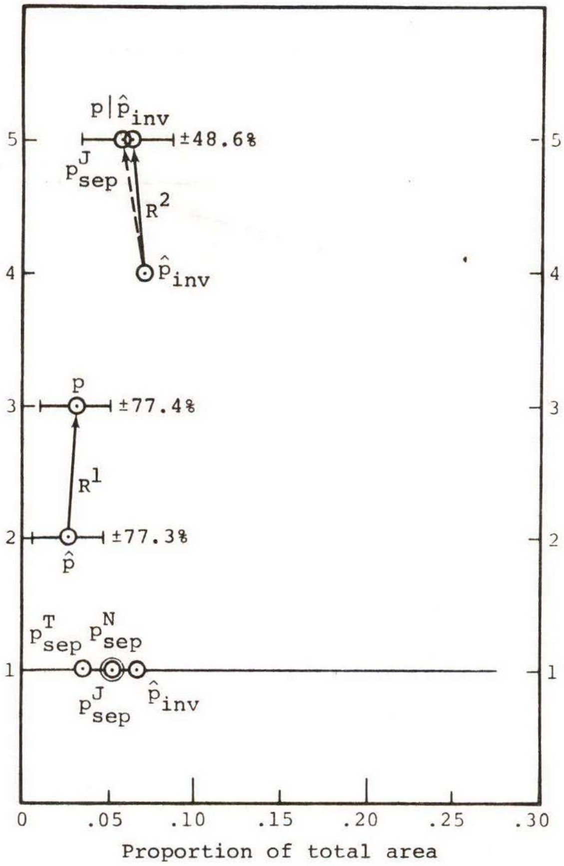
[a]Forest = softwood plus hardwood.

[b]Not classified.

Table key

R^1 = regression transformation

R^2 = applied regression



<u>Line number</u>

1 — ADP inventory ($\hat{p}_{inv}$) and separability estimates (p_{sep})

2 — simulated inventory estimate from 10 samples ($\hat{p}$)

3 — aerial photointerpretation estimate from 10 samples (p)

4 — simulated inventory ($\hat{p}_{inv}$)

5 — regression estimate

<u>Legend</u>

R^1 — regression transformation

R^2 — applied regression

J — July

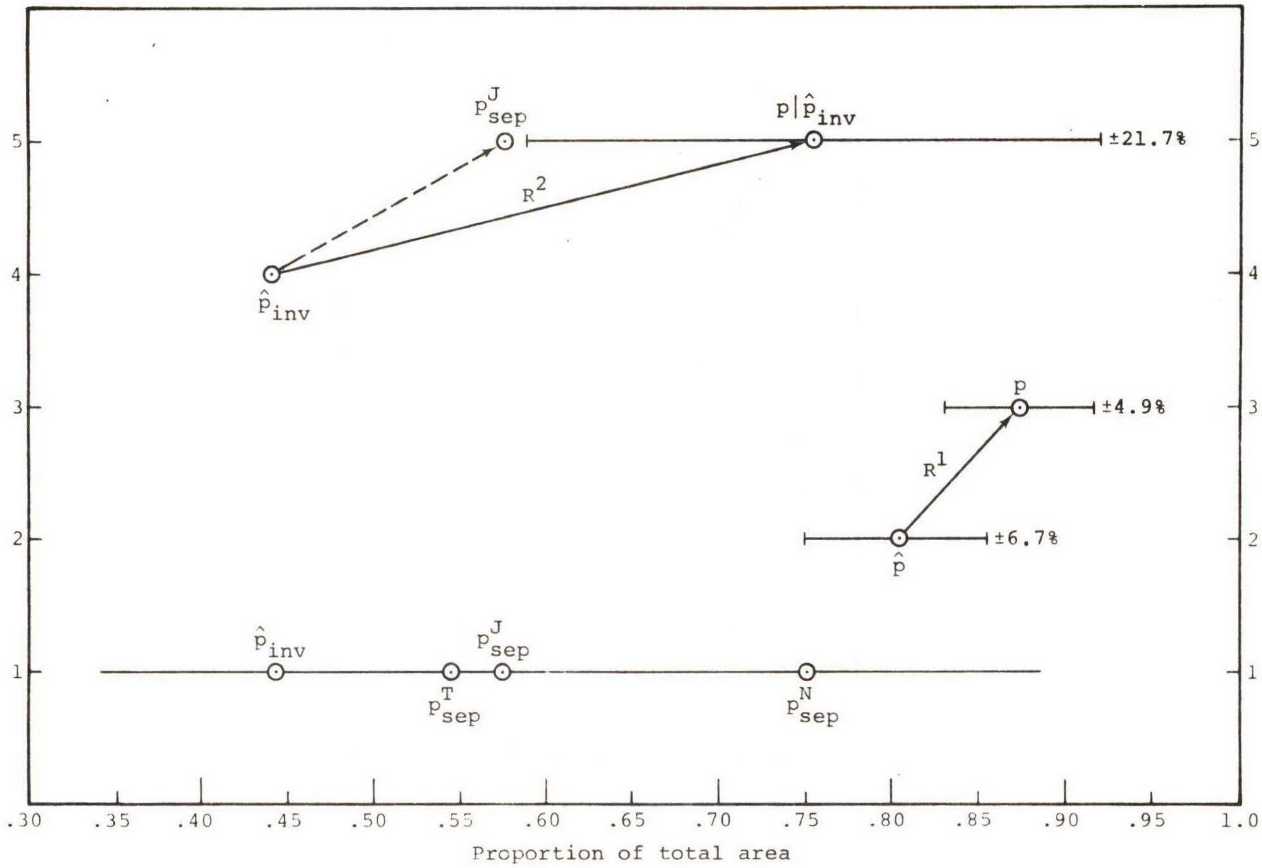
N — November

T — temporal

⟶ — actual transformation

--➝ — expected transformation

Figure C-1.— Proportion estimates of softwood.



<u>Line number</u>

1 — ADP inventory ($\hat{p}_{inv}$) and separability estimates (p_{sep})

2 — simulated inventory estimate from 10 samples ($\hat{p}$)

3 — aerial photointerpretation estimate from 10 samples (p)

4 — simulated inventory ($\hat{p}_{inv}$)

5 — regression estimate

<u>Legend</u>

R^1 — regression transformation

R^2 — applied regression

J — July

N — November

T — temporal

→ — actual transformation

- - → — expected transformation

Figure C-2.— Proportion estimates of hardwood.

APPENDIX D

REFERENCES

APPENDIX D

REFERENCES

1. Kan, E. P.: Ten-Ecosystem Study (TES) Investigation Plan. Lockheed Electronics Co., Inc. (Houston), LEC-8667, July 1976.

2. Kan, E. P., ed.: Technical Analysis Procedures for the Ten-Ecosystem Study. Lockheed Electronics Co., Inc. (Houston), LEC-9379, Dec. 1976.

3. Luedke, William J.: Ayer Directory of Publications. Ayer Press (Philadelphia), 1978.

4. Essex, Burton L.; and Spencer, John S., Jr.: Timber Resource of Missouri's Eastern Ozarks, 1972. North Central Forest Experiment Station, USDA Forest Service (St. Paul, Minn.), 1974.

5. Plan for Managing the National Forests in Missouri: Mark Twain and Clark. Eastern Region, USDA Forest Service, (Rolla, Mo.), 1975.

6. Oetking, P.; Feray, D. E.; and Renfro, H. B.: Geological Highway Map of the Mid-Continent Region. American Assoc. Petrol. Geologists (Tulsa), 1966.

7. Dillman, R. D.; and Ward, J. F.: Report on Site Selection for the Ten-Ecosystem Study (TES). Lockheed Electronics Co., Inc. (Houston), LEC-8952, Oct. 1976.

8. Scrivner, C. L.; Baker, J. C.; and Miller, B. J.: Soils of Missouri: A Guide to Their Identification and Interpretation. Extension Division, Univ. of Mo., Jan. 1975.

9. Martin, V. J.; et al.: Missouri Soil and Water Conservation Needs Inventory. USDA Soil Conservation Service (Columbia, Mo.), 1970.